

# Sublinear Estimation of Weighted Matchings in Dynamic Data Streams\*

Marc Bury, Chris Schwiegelshohn

*Efficient Algorithms and Complexity Theory, TU Dortmund, Germany*  
*{firstname.lastname}@tu-dortmund.de*

This paper presents an algorithm for estimating the weight of a maximum weighted matching by augmenting any estimation routine for the size of an unweighted matching. The algorithm is implementable in any streaming model including dynamic graph streams. We also give the first constant estimation for the maximum matching size in a dynamic graph stream for planar graphs (or any graph with bounded arboricity) using  $\tilde{O}(n^{4/5})$  space which also extends to weighted matching. Using previous results by Kapralov, Khanna, and Sudan (2014) we obtain a  $\text{polylog}(n)$  approximation for general graphs using  $\text{polylog}(n)$  space in random order streams, respectively. In addition, we give a space lower bound of  $\Omega(n^{1-\varepsilon})$  for any randomized algorithm estimating the size of a maximum matching up to a  $1 + O(\varepsilon)$  factor for adversarial streams.

## 1 Introduction

Large graph structures encountered in social networks or the web-graph have become focus of analysis both from theory and practice. To process such large input, conventional algorithms often require an infeasible amount of running time, space or both, giving rise to other models of computation. Much theoretical research focuses on the streaming model where the input arrives one by one with the goal of storing as much information as possible in small, preferably polylogarithmic, space. Streaming algorithms on graphs were first studied by Henzinger et al. [19], who showed that even simple problems often admit no solution with such small space requirements. The semi-streaming model [16] where the stream consists of the edges of a graph and the algorithm is allowed  $O(n \cdot \text{polylog}(n))$  space and allows few (ideally just one) passes over the data relaxes these requirements and has received considerable attention. Problems studied in the semi-streaming model include sparsification, spanners, connectivity, minimum spanning trees, counting triangles and matching, for an overview we refer to a recent survey by McGregor [30]. Due to the fact that graphs motivating this research are dynamic structures that change over time there has recently been research on streaming algorithms supporting deletions. We now review the literature on streaming algorithms for matching and dynamic streams.

---

\*Supported by Deutsche Forschungsgemeinschaft, grant BO 2755/1-2 and within the Collaborative Research Center SFB 876, project A2

**Matching** Maintaining a 2 approximation to the maximum matching (MM) in an insertion-only stream can be straightforwardly done by greedily maintaining a maximal matching [16]. Improving on this algorithm turns out to be difficult as Goel et al. [18] showed that no algorithm using  $\tilde{O}(n)$  space can achieve an approximation ratio better than  $\frac{3}{2}$  which was improved by Kapralov to  $\frac{e}{e-1}$  [22]. Konrad et al. [25] gave an algorithm using  $\tilde{O}(n)$  space with an approximation factor of 1.989 if the edges are assumed to arrive in random order. For weighted matching (MWM), a series of results have been published [16, 29, 13, 34, 14] with the current best bound of  $4 + \varepsilon$  being due to Crouch and Stubbs [11].

To bypass the natural  $\Omega(n)$  bound required by any algorithm maintaining an approximate matching, recent research has begun to focus on estimating the size of the maximum matching. Kapralov et al. [23] gave a polylogarithmic approximate estimate using polylogarithmic space for random order streams. For certain sparse graphs including planar graphs, Esfandiari et al. [15] describe how to obtain a constant factor estimation using  $\tilde{O}(n^{2/3})$  space in a single pass and  $\tilde{O}(\sqrt{n})$  space using two passes or assuming randomly ordered streams. The authors also gave a lower bound of  $\Omega(\sqrt{n})$  for any approximation better than  $\frac{3}{2}$ .

**Dynamic Streams** In the turnstile model, the stream consists of a sequence of additive updates to a vector. Problems studied in this model include numerical linear algebra problems such as regression and low-rank approximation, and maintaining certain statistics of a vector like frequency moments, heavy hitters or entropy. Linear sketches have proven to be the algorithmic technique of choice and might as well be the only algorithmic tool able to efficiently do so, see Li, Nguyen and Woodruff [27]. Dynamic graphs as introduced and studied by Ahn, Guha and McGregor [1, 2, 3, 4] are similar to, but weaker than turnstile updates. Though both streaming models assume update to the input matrix, there usually exists a consistency assumption for streams, i.e. at any given time the multiplicity of an edge is either 0 or 1 and edge weights cannot change arbitrarily but are first set to 0 and then reinserted with the desired weight. The authors extend some of the aforementioned problems such as connectivity, sparsification and minimum spanning trees to this setting. Recent results by Assadi et al. [5] showed that approximating matchings in dynamic streams is hard by providing a space lower bound of  $\Omega(n^{2-3\varepsilon})$  for approximating the maximum matching within a factor of  $\tilde{O}(n^\varepsilon)$ . Simultaneously, Konrad [24] showed a similar but slightly weaker lower bound of  $\Omega(n^{3/2-4\varepsilon})$ . Both works presented an algorithm with an almost matching upper bound on the space complexity of  $\tilde{O}(n^{2-2\varepsilon})$  [24] and  $\tilde{O}(n^{2-3\varepsilon})$  [5]. Chitnis et al. [8] gave a streaming algorithm using  $\tilde{O}(k^2)$  space that returns an exact maximum matching under the assumption that the size is at most  $k$ . It is important to note that all these results actually compute a matching. In terms of estimating the size of the maximum matching, Chitnis et al. [8] extended the estimation algorithms for sparse graphs from [15] to the settings of dynamic streams using  $\tilde{O}(n^{4/5})$  space. A bridge between dynamic graphs and the insertion-only streaming model is the sliding window model studied by Crouch et al. [10]. The authors give a  $(3 + \varepsilon)$ -approximation algorithm for maximum matching.

The  $p$ -Schatten norm of a matrix  $A$  is defined as the  $\ell_p$ -norm of the vector of singular values. It is well known that computing the maximum matching size is equivalent to computing the rank of the Tutte matrix [31, 28] (see also Section 2.1). Estimating the maximum matching size therefore is a special case of estimating the rank or 0-Schatten norm of a matrix. Li, Nguyen and Woodruff gave strong lower bounds on the space requirement for estimating Schatten norms in dynamic streams [26]. Any estimation of the rank within any constant factor is shown to require  $\Omega(n^2)$  space when using bi-linear sketches and  $\Omega(\sqrt{n})$  space for general linear sketches.

## Techniques and Contribution

Table 1 gives an overview of our results in comparison to previously known algorithms and lower bounds. Our first main result (Section 2) is an approximate estimation algorithm for the maximum

|             | Reference | Graph class        | Streaming model | Approx. factor              | Space                               |
|-------------|-----------|--------------------|-----------------|-----------------------------|-------------------------------------|
| <b>MM:</b>  | Greedy    | General            | Adversarial     | 2                           | $O(n)$                              |
|             | [23]      | General            | Random          | $\text{polylog}(n)$         | $\text{polylog}(n)$                 |
|             | [15]      | Trees              | Adversarial     | $2 + \varepsilon$           | $\tilde{O}(\sqrt{n})$               |
|             | [15]      | Bounded arboricity | Adversarial     | $O(1)$                      | $\tilde{O}(n^{2/3})$                |
|             | here      | Trees              | Dynamic         | $2 + \varepsilon$           | $O(\frac{\log^2 n}{\varepsilon^2})$ |
|             | here      | Bounded arboricity | Dynamic         | $O(1)$                      | $\tilde{O}(n^{4/5})$                |
|             | [15]      | Forests            | Adversarial     | $\frac{3}{2} - \varepsilon$ | $\Omega(\sqrt{n})$                  |
| <b>MWM:</b> | here      | General            | Adversarial     | $1 + O(\varepsilon)$        | $\Omega(n^{1-\varepsilon})$         |
|             | [11]      | General            | Adversarial     | $4 + \varepsilon$           | $O(n \log^2 n)$                     |
|             | here      | General            | Random          | $\text{polylog}(n)$         | $\text{polylog}(n)$                 |
|             | here      | Bounded arboricity | Dynamic         | $O(1)$                      | $\tilde{O}(n^{4/5})$                |

Table 1: Results for estimating the size (weight) of a maximum (weighted) matching in data streams.

weight of a matching. We give a generic procedure using any unweighted estimation as black box. In particular:

**Theorem 1** (informal version). *Given a  $\lambda$ -approximate estimation using  $S$  space, there exists an  $O(\lambda^4)$ -approximate estimation algorithm for the weighted matching problem using  $O(S \cdot \log n)$  space.*

The previous algorithms for weighted matchings in insertion only streams analyzed in [16, 29, 13, 34] extend the greedy approach by a charging scheme. If edges are mutually exclusive, the new edge will be added if the weight of the matching increases by a given threshold, implicitly partitioning the edges into sets of geometrically increasing weights. We use a similar scheme, but with a twist: Single edge weights cannot be charged to an edge with larger weight as estimation routines do not necessarily give information on distinct edges. However, entire matchings can be charged as the contribution of a specific range of weights  $r$  can only be large if these edges take up a significant part of any maximum matching in the subgraph containing only the edges of weight at least  $r$ . For analysis, we use a result on parallel algorithms by Uehara and Chen [32]. We show that the weight outputted by our algorithm is close to the weight of the matching computed by the authors, implying an approximation to the maximum weight.

We can implement this algorithm in dynamic streams although at submission, we were unaware of any estimations for dynamic streams. Building on the work by Esfandiari et al. [15], we give a constant estimation on the matching size in bounded arboricity graphs. The main obstacle to adapt their algorithms for bounded arboricity graphs is that they maintain a small size matching using the greedy algorithm which is hard for dynamic streams. Instead of maintaining a matching, we use the Tutte matrix to get a 1-pass streaming algorithm using  $\tilde{O}(n^{4/5})$  space, which immediately extends to weighted matching. Similar bounds have been obtained independently by Chitnis et al. [8].

Our lower bound (Section 3) is proven via reduction from the Boolean Hidden Hypermatching problem introduced by Verbin and Yu [33]. In this setting, two players Alice and Bob are given a binary  $n$ -bit string and a perfect  $t$ -hypermatching on  $n$  nodes, respectively. Bob also gets a binary string  $w$ . The players are promised that the parity of bits corresponding to the nodes of the  $i$ -th hypermatching either are equal to  $w_i$  for all  $i$  or equal to  $1 - w_i$  for all  $i$  and the task is to find out which case holds using only a single round of communication. We construct a graph consisting of a  $t$ -clique for each hyperedge of Bob's matching and a single edge for each bit of Alice's input that has one node in common with the  $t$ -cliques. Then we show that approximating the matching size

---

**Algorithm 1** Approximation of Weighted Matching from [32]

---

**Require:** Graph  $G = (V, E = \bigcup_{i=1}^t E_i)$

**Ensure:** Matching

**for**  $i = t$  to 1 **do**

    Find a maximal matching  $M_i$  in  $G_i = (V, E_i)$ .

    Remove all edges  $e$  from  $E$  such that  $e \in M_i$  or  $e$  shares a node with an edge in  $M_i$ .

**return**  $\bigcup_{i=1}^t M_i$ 

---

within a factor better than  $1 + O(1/t)$  can also solve the Boolean Hidden Hypermatching instance. Using the lower bound of  $\Omega(n^{1-1/t})$  from [33] we have

**Theorem 2** (informal version). *Any 1-pass streaming algorithm approximating the size of the maximum matching matching up to an  $(1 + O(\varepsilon))$  factor requires  $\Omega(n^{1-\varepsilon})$  bits of space.*

This lower bound also implies an  $\Omega(n^{1-\varepsilon})$  space bound for  $1 + O(\varepsilon)$  approximating the rank of a matrix in data streams which also improves the  $\Omega(\sqrt{n})$  bound by Li, Nguyen, and Woodruff [26] for linear sketches.

## 1.1 Preliminaries

We use  $\tilde{O}(f(n))$  to hide factors polylogarithmic in  $f(n)$ . Any randomized algorithm succeeding with high probability has at least  $1 - 1/n$  chance of success. Graphs are denoted by  $G(V, E, w)$  where  $V$  is the set of  $n$  nodes,  $E$  is the set of edges and  $w : E \rightarrow \mathbb{R}^+$  is a weight function. Our estimated value  $\widehat{M}$  is a  $\lambda$ -approximation to the size of the maximum matching  $M$  if  $\widehat{M} \leq |M| \leq \lambda \widehat{M}$ .

## 2 Weighted Matching

We start by describing the parallel algorithm by Uehara and Chen [32], see Algorithm 1. Let  $\gamma > 1$  and  $k > 0$  be constant. We partition the edge set by  $t$  ranks where all edges  $e$  in rank  $i \in \{1, \dots, t\}$  have a weight  $w(e) \in (\gamma^{i-1} \cdot \frac{w_{max}}{kN}, \gamma^i \cdot \frac{w_{max}}{kN}]$  where  $w_{max}$  is the maximal weight in  $G$ . Let  $G' = (V, E, w)$  be equal to  $G$  but each edge  $e$  in rank  $i$  has weight  $r_i := \gamma^i$  for all  $i = 1, \dots, t$ . Starting with  $i = t$ , we compute an unweighted maximal matching  $M_i$  considering only edges in rank  $i$  (in  $G'$ ) and remove all edges incident to a matched node. Continue with  $i - 1$ . The weight of the matching  $M = \bigcup M_i$  is  $w(M) = \sum_{i=1}^t r_i \cdot |M_i|$  and satisfies  $w_G(M^*) \geq w_{G'}(M) \geq \frac{1}{2\gamma} \cdot w_G(M^*)$  where  $M^*$  is an optimal weighted matching in  $G$ . The previous algorithms [16, 29, 13, 34, 11] for insertion-only streams use a similar partitioning of edge weights. Since these algorithms are limited to storing one maximal matching (in case of [11] one maximal matching per rank), they cannot compute residual maximal matchings in each rank. However, by charging the smaller edge weights into the higher ones, the resulting approximation factor can be made reasonably close to that of Uehara and Chen. Since these algorithms maintain matchings, they cannot have sublinear space in an insertion-only stream and they need at least  $\Omega(n^{2-3\varepsilon})$  in a dynamic stream even when the maintained matching is only a  $O(n^\varepsilon)$  approximation ([5]). Though the complexity for unweighted estimating unweighted matchings is not settled for any streaming model, there exist graph classes for which one can improve on these algorithms wrt space requirement. Therefore, we assume the existence of a black box  $\lambda$ -approximate matching estimation algorithm.

### Algorithm and Analysis

In order to adapt this idea to our setting, we need to work out the key properties of the partitioning and how we can implement it in a stream. The first problem is that we cannot know  $w_{max}$  in

a stream a priori and in a dynamic stream even maintaining  $w_{max}$  is difficult. However, the appropriate partition of an inserted edge depends on  $w_{max}$ . Recalling the partitioning of Uehara and Chen, we disregard all edges with weight smaller than  $\frac{w_{max}}{kN}$  which is possible because the contribution of these edges is at most  $\frac{N}{2} \cdot \frac{w_{max}}{kN} = \frac{w_{max}}{2k} \leq \frac{OPT}{2k}$  where  $OPT$  is the weight of an optimal weighted matching. Thus, we can only consider edges with larger weight and it is also possible to partition the set of edges in a logarithmic number of sets. Here, we use the properties that edge weights within a single partition set are similar and that  $\frac{1}{\gamma} \leq \frac{w(e)}{w(e')} \leq \gamma$  for two edges  $e \in E_i$  and  $e' \in E_{i-1}$  with  $i \in \{2, \dots, t\}$ . These properties are sufficient to get a good approximation on the optimal weighted matching which we show in the next lemma. The proof is essentially the same as in [32].

**Lemma 1.** *Let  $G = (V, E, w)$  be a weighted graph and  $\varepsilon > 0$  be an approximation parameter. If a partitioning  $E_1, \dots, E_t$  of  $E$  and a weight function  $w' : E \rightarrow \mathbb{R}$  satisfy*

$$\frac{1}{1+\varepsilon} \leq \frac{w'(e)}{w(e)} \leq 1 \text{ for all } e \in E \quad \text{and} \quad \frac{w(e_1)}{w(e_2)} \leq 1+\varepsilon \quad \text{and} \quad w(e) < w(e')$$

*for all choices of edges  $e_1, e_2 \in E_i$  and  $e \in E_i, e' \in E_j$  with  $i < j$  and  $i, j \in \{1, \dots, t\}$  then Algorithm 1 returns a matching  $M = \bigcup_{i=1}^t M_i$  with*

$$\frac{1}{2(1+\varepsilon)^2} \cdot w(M^*) \leq w'(M) \leq w(M^*)$$

*where  $M^*$  is an optimal weighted matching in  $G$ .*

*Proof.* The first property  $\frac{1}{1+\varepsilon} \leq \frac{w'(e)}{w(e)} \leq 1$  for all  $e \in E$  implies that  $\frac{w(S)}{1+\varepsilon} \leq w'(S) \leq w(S)$  for every set of edges  $S \subseteq E$ . Thus, it remains to show that  $\frac{1}{2(1+\varepsilon)} \cdot w(M^*) \leq w(M) \leq w(M^*)$ . Since  $M^*$  is an optimal weighted matching, it is clear that  $w(M) \leq w(M^*)$ . For the lower bound, we distribute the weight of the edges from the optimal solution to edges in  $M$ . Let  $e \in M^*$  and  $i \in \{1, \dots, t\}$  such that  $e \in E_i$ . We consider the following cases:

1.  $e \in M_i$ : We charge the weight  $w(e)$  to the edge itself.
2.  $e \notin M_i$  but at least one node incident to  $e$  is matched by an edge in  $M_i$ : Let  $e' \in M_i$  be an edge sharing a node with  $e$ . Distribute the weight  $w(e)$  to  $e'$ .
3.  $e \notin M_i$  and there is no edge in  $M_i$  sharing a node with  $e$ : By Algorithm 1, there has to be an edge  $e' \in M_j$  with  $j > i$  which shares a node with  $e$ . We distribute the weight  $w(e)$  to  $e'$ .

Since  $M^*$  is a matching, there can only be at most two edges from  $M^*$  distributing their weights to an edge in  $M$ . We know that  $\frac{w(e)}{w(e')} \leq 1+\varepsilon$  for all choices of two edges  $e, e' \in E_i$  with  $i \in \{1, \dots, t\}$  which means that in the case 2. we have  $w(e) \leq (1+\varepsilon) \cdot w(e')$ . In case 3. it holds  $w(e) < w(e')$ . Thus, the weight distributed to an edge  $e'$  in  $M$  is at most  $2(1+\varepsilon)w(e')$ . This implies that  $w(M^*) = \sum_{e \in M^*} w(e) \leq \sum_{e' \in M} 2(1+\varepsilon) \cdot w(e') = 2(1+\varepsilon) \cdot w(M)$  which concludes the proof.  $\square$

Using Lemma 1, we can partition the edge set in a stream in an almost oblivious manner: Let  $(e_0, w(e_0))$  be the first inserted edge. Then an edge  $e$  belongs to  $E_i$  iff  $2^{i-1} \cdot w(e_0) < w(e) \leq 2^i \cdot w(e_0)$  for some  $i \in \mathbb{Z}$ . For the sake of simplicity, we assume that the edge weights are in  $[1, W]$ . Then the number of sets is  $\mathcal{O}(\log W)$ . We would typically expect  $W \in \text{poly } n$  as otherwise storing weights becomes infeasible.

We now introduce a bit of notation we will use in the algorithm and throughout the proof. We partition the edge set  $E = \bigcup_{i=0}^t E_i$  by  $t+1 = \mathcal{O}(\log W)$  ranks where the set  $E_i$  contains all edges  $e$  with weight  $w(e) \in [2^i, 2^{i+1})$ . Wlog we assume  $E_t \neq \emptyset$  (otherwise let  $t$  be the largest rank with

---

**Algorithm 2** Weighted Matching Approximation

---

**Require:** Graph  $G = (V, \bigcup_{i=0}^t E_i)$  with weights  $r_i$  for edges in  $E_i$

**Ensure:** Estimator of the weighted matching

```
for  $i = t$  to 0 do
     $\widehat{S}_i = \widehat{R}_i = 0$ 
weight = 0, last = t
 $\widehat{R}_t = \widehat{S}_t = \text{Unweighted Matching Estimation}(V, E_t)$ 
for  $i = t - 1$  to 0 do
     $\widehat{S}_i = \text{Unweighted Matching Estimation}(V, \bigcup_{j=i}^t E_j)$ 
    if  $\widehat{S}_i > \widehat{S}_{last} \cdot T$  then ▷ Add current index  $i$  to  $I_{good}$ 
        if  $\widehat{S}_i - \widehat{S}_{last} \geq c \cdot \widehat{R}_{last}$  then ▷ Add current index  $i$  to  $I_{sign}$ 
             $\widehat{R}_i = \widehat{S}_i - \widehat{S}_{last}$ 
            last = i
        else
             $\widehat{S}_i = 0$ 
return  $\frac{2}{5} \sum_{i=0}^t r_i \cdot \widehat{R}_i$ 
```

---

$E_t \neq \emptyset$ ). Let  $G' = (V, E, w')$  be equal to  $G$  but each edge  $e \in E_i$  has weight  $w'(e) = r_i := 2^i$  for all  $i = 0, \dots, t$ . Let  $M = \bigcup_{i=0}^t M_i$  be the matching computed by the partitioning algorithm and  $S$  be a  $(t+1)$ -dimensional vector with  $S_i = \sum_{j=i}^t |M_j|$ .

Algorithm 2 now proceeds as follows: For every  $i \in \{0, \dots, t\}$  the size of a maximum matching in  $(V, \bigcup_{j=i}^t E_j)$  and  $S_i$  differ by only a constant factor. Conceptually, we set our estimator  $\widehat{S}_i$  of  $S_i$  to be the approximation of the size of the maximum matching of  $(V, \bigcup_{j=i}^t E_i)$  and the estimator of the contribution of the edges in  $E_i$  to the weight of an optimal weighted matching is  $\widehat{R}_i = \widehat{S}_i - \widehat{S}_{i+1}$ . The estimator  $\widehat{R}_i$  is crude and generally not a good approximation to  $|M_i|$ . What helps us is that if the edges  $M_i$  have a significant contribution to  $w(M)$ , then  $|M_i| \gg \sum_{j=i+1}^t |M_j| = S_{i+1}$ . In order to detect whether the matching  $M_i$  has a significant contribution to the objective value, we introduce two parameters  $T$  and  $c$ . The first matching  $M_t$  is always significant (and the simplest to approximate by setting  $\widehat{R}_t = \widehat{S}_t$ ). For all subsequent matchings  $i < t$ , let  $j$  be the most recent matching which we deemed to be significant. We require  $\widehat{S}_i \geq T \cdot \widehat{S}_j$  and  $\widehat{R}_i \geq c \cdot \widehat{R}_j$ . If both criteria are satisfied, we use the estimator  $\widehat{R}_i = \widehat{S}_i - \widehat{S}_j$  and set  $i$  to be the now most recent, significant matching, otherwise we set  $\widehat{R}_i = 0$ . The final estimator of the weight is  $\sum_{i=0}^t r_i \cdot \widehat{R}_i$ . The next definition gives a more detailed description of the two sets of ranks which are important for the analysis.

**Definition 1** (Good and Significant Ranks). Let  $\widehat{S}$  and  $\widehat{R}$  be the vectors at the end of Algorithm 2. An index  $i$  is called to be a *good rank* if  $\widehat{S}_i \neq 0$  and  $i$  is a *significant rank* if  $\widehat{R}_i \neq 0$ . We denote the set of good ranks by  $I_{good}$  and the set of significant ranks by  $I_{sign}$ , i.e.,  $I_{good} := \{i \subseteq \{0, \dots, t\} \mid \widehat{S}_i \neq 0\}$  and  $I_{sign} := \{i \subseteq \{0, \dots, t\} \mid \widehat{R}_i \neq 0\}$ . We define  $I_{good}$  and  $I_{sign}$  to be in descending order and we will refer to the  $\ell$ -th element of  $I_{good}$  and  $I_{sign}$  by  $I_{good}(\ell)$  and  $I_{sign}(\ell)$ , respectively. That means  $I_{good}(1) > I_{good}(2) > \dots > I_{good}(|I_{good}|)$  and  $I_{sign}(1) > I_{sign}(2) > \dots > I_{sign}(|I_{sign}|)$ . We slightly abuse the notation and set  $I_{sign}(|I_{sign}| + 1) = 0$ . Let  $D_1 := |M_t|$  and for  $\ell \in \{2, \dots, |I_{sign}|\}$  we define the sum of the matching sizes between two significant ranks  $I_{sign}(\ell)$  and  $I_{sign}(\ell - 1)$  where the smaller significant rank is included by  $D_\ell := \sum_{i=I_{sign}(\ell)}^{I_{sign}(\ell-1)-1} |M_i|$ .

In the following, we subscript indices by  $s$  for significant ranks and by  $g$  for good ranks for the

sake of readability. Looking at Algorithm 2 we can proof some simple properties of  $I_{good}$  and  $I_{sign}$ .

**Lemma 2.** *Let  $I_{good}$  and  $I_{sign}$  be defined as in Definition 1. Then*

1.  $I_{good}(1) = I_{sign}(1) = t$  and  $I_{sign} \subseteq I_{good}$ .
2. For every good rank  $i_g \in I_{good}$  there is an  $\ell \in \{0, \dots, |I_{sign}|\}$  such that  $I_{sign}(\ell) > i_g \geq I_{sign}(\ell + 1)$  and  $\widehat{S}_{i_g} > T \cdot \widehat{S}_{I_{sign}(\ell)}$ .
3. For every  $i_s, i'_s \in I_{sign}$  with  $i_s < i'_s$  it holds  $\widehat{S}_{i_s} > T \cdot \widehat{S}_{i'_s}$ .
4. For any  $i_s \in I_{sign}$  and  $i'_s \in I_{sign}$  with  $i'_s < i_s$  it is  $\widehat{R}_{i'_s} > c \cdot \widehat{R}_{i_s}$ .
5. For any  $i_s \in I_{sign}$  and  $i_g \in I_{good}$  with  $i_g < i_s$  it is  $\widehat{S}_{i_g} > T \cdot \widehat{S}_{i_s}$ .

*Proof.*

1. It is clear that  $I_{sign} \subseteq I_{good}$ . Since we assumed that  $E_t \neq \emptyset$ , there is a nonempty matching in  $E_t$  which means that  $\widehat{S}_t = \widehat{R}_t > 0$ .
2. Let  $\ell$  be the position of *last* in  $I_{sign}$  where *last* is the value of the variable in Algorithm 2 during the iteration  $i = i_g$ . Then  $I_{sign}(\ell) > i_g \geq I_{sign}(\ell + 1)$  (recall that we defined  $I_{sign}(|I_{sign}| + 1) = 0$ ). Since  $i_g$  is good, it is  $\widehat{S}_{i_g} > T \cdot \widehat{S}_{last} = \widehat{S}_{I_{sign}(\ell)}$ .
3. Since significant ranks are also good, we can apply 2. to get  $\widehat{S}_{I_{sign}(\ell+1)} > T \cdot \widehat{S}_{I_{sign}(\ell)}$  where  $I_{sign}(\ell + 1) < I_{sign}(\ell)$ . By transitivity this implies the statement.
4. For every  $i_s \in I_{sign}$  we have  $\widehat{R}_{i_s} \geq c \cdot \widehat{R}_{last}$  where *last* is the value of the variable in Algorithm 2 in iteration  $i = i_s$ . By definition it is  $last \in I_{sign}$  and  $last > i_s$ . Therefore, it holds  $\widehat{R}_{I_{sign}(\ell+1)} > c \cdot \widehat{R}_{I_{sign}(\ell)}$  for every  $\ell \in \{0, \dots, |I_{sign}| - 1\}$  which implies the statement.
5. Using 2. we know that  $\widehat{S}_{i_g} > T \cdot \widehat{S}_{I_{sign}(\ell)}$  for some  $\ell \in \{0, \dots, |I_{sign}|\}$ . If  $i_s$  is equal to  $I_{sign}(\ell)$  then we are done. Otherwise, we have  $i_s > I_{sign}(\ell)$  and we can use 3. to get  $\widehat{S}_{i_g} > T \cdot \widehat{S}_{I_{sign}(\ell)} > T \cdot \widehat{S}_{i_s}$ .

□

Now, we have the necessary notations and properties of good and significant ranks to proof our main theorem.

**Theorem 1.** *Let  $G = (V, E, w)$  be a weighted graph where the weights are from  $[1, W]$ . Let  $A$  be an algorithm that returns an  $\lambda$ -estimator  $\widehat{M}$  for the size of a maximum matching  $M$  of a graph with  $1/\lambda \cdot |M| \leq \widehat{M} \leq |M|$  with failure probability at most  $\delta$  and needs space  $S$ . If we partition the edge set into sets  $E_0, \dots, E_t$  with  $t = \lfloor \log W \rfloor$  where  $E_i$  consists of all edges with weight in  $[2^i, 2^{i+1})$ , set  $r_i = 2^i$ , and use  $A$  as the unweighted matching estimator in Algorithm 2, then there are parameters  $T$  and  $c$  depending on  $\lambda$  such that the algorithm returns an  $O(\lambda^4)$ -estimator  $\widehat{W}$  for the weight of the maximum weighted matching with failure probability at most  $\delta \cdot (t + 1)$  using  $O(S \cdot t)$  space, i.e. there is a constant  $c$  such that  $\frac{1}{c\lambda^4} \cdot w(M^*) \leq \widehat{W} \leq w(M^*)$  where  $M^*$  is an optimal weighted matching.*

*Proof.* In the following we condition on the event that all calls to the unweighted estimation routine succeed, which happens with probability at least  $1 - \delta \cdot (t+1)$ . The estimator returned by Algorithm 2 can be written as  $\sum_{\ell=1}^{|I_{\text{sign}}|} r_{I_{\text{sign}}(\ell)} \cdot \widehat{R_{I_{\text{sign}}(\ell)}}$ . Using similar arguments as found in Lemma 4 of [32], we have  $\frac{1}{8} \cdot w(M^*) \leq \sum_{i=0}^t r_i |M_i| \leq w(M^*)$ . Thus, it is sufficient to show that  $\sum_{\ell=1}^{|I_{\text{sign}}|} r_{I_{\text{sign}}(\ell)} \cdot \widehat{R_{I_{\text{sign}}(\ell)}}$  is a good estimator for  $\sum_{i=0}^t r_i |M_i|$ . We first consider the problem of estimating  $D_\ell$ , and then how to charge the matching sizes.

### (1) Estimation of $D_\ell$

Since  $\bigcup_{j=i}^t M_j$  is a maximal matching in  $\bigcup_{j=i}^t E_j$ ,  $\widehat{S}_i$  is a good estimator for  $S_i$ :

**Lemma 3.** *For all  $i \in \{0, \dots, t\}$  we have  $\frac{1}{\lambda} \cdot S_i \leq \widehat{S}_i \leq 2 \cdot S_i$ .*

*Proof.* Let  $F_j$  be the set of unmatched nodes after the iteration  $j$  of Algorithm 1. Let  $M^*$  be a maximum matching in  $(V, \bigcup_{j=i}^t E_j)$ .  $M_j$  is a maximal matching of  $(V, E_j(F_j))$  and therefore  $\bigcup_{j=i}^t M_j$  is a maximal matching of  $(V, \bigcup_{j=i}^t E_j)$ . This allows us to apply the bounds of the  $\lambda$ -approximate estimation algorithm:

$$\frac{1}{\lambda} \cdot S_i = \frac{1}{\lambda} \cdot \sum_{j=i}^t |M_j| \leq \frac{1}{\lambda} \cdot |M^*| \leq |\widehat{S}_i| \leq |M^*| \leq 2 \cdot \sum_{j=i}^t |M_j| = 2 \cdot S_i.$$

□

Next, we show that for an index  $i_g \in I_{\text{good}}$  the difference  $\widehat{S}_{i_g} - \widehat{S}_{I_{\text{sign}}(\ell)}$  to the last significant rank is a good estimator for  $\sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i|$ .

**Lemma 4.** *For all  $i_g \in I_{\text{good}}$  with  $I_{\text{sign}}(\ell+1) \leq i_g < I_{\text{sign}}(\ell)$  for some  $\ell \in \{1, \dots, |I_{\text{sign}}|\}$  and  $T = 8\lambda^2 - 2\lambda$ ,*

$$\frac{1}{2\lambda} \cdot \sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i| < \widehat{S}_{i_g} - \widehat{S}_{I_{\text{sign}}(\ell)} < \frac{5}{2} \cdot \sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i|$$

and  $\frac{1}{\lambda} |M_t| \leq \widehat{S}_t \leq 2 |M_t|$ .

*Proof.* For all  $i_g \in I_{\text{good}}$  with  $I_{\text{sign}}(\ell+1) \leq i_g < I_{\text{sign}}(\ell)$  we have

$$\begin{aligned} \sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i| &= S_{i_g} - S_{I_{\text{sign}}(\ell)} \stackrel{\text{Lem. 3}}{\geq} \frac{1}{2} \cdot \widehat{S}_{i_g} - \lambda \cdot \widehat{S}_{I_{\text{sign}}(\ell)} \\ &\stackrel{\text{Lem. 2 (2)}}{>} \frac{T}{2} \cdot \widehat{S}_{I_{\text{sign}}(\ell)} - \lambda \cdot \widehat{S}_{I_{\text{sign}}(\ell)} \stackrel{\text{Lem. 3}}{\geq} \left( \frac{T}{2} - \lambda \right) \cdot \frac{1}{\lambda} \cdot S_{I_{\text{sign}}(\ell)} \\ &= \frac{T - 2\lambda}{2\lambda} \cdot S_{I_{\text{sign}}(\ell)}, \end{aligned} \tag{1}$$

Setting  $T = 8\lambda^2 - 2\lambda$ , we then obtain the following upper and lower bounds

$$\begin{aligned} \widehat{S}_{i_g} - \widehat{S}_{I_{\text{sign}}(\ell)} &\stackrel{\text{Lem. 3}}{\geq} \frac{1}{\lambda} \cdot S_{i_g} - 2 \cdot S_{I_{\text{sign}}(\ell)} = \frac{1}{\lambda} \sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i| - \left( 2 - \frac{1}{\lambda} \right) \cdot S_{I_{\text{sign}}(\ell)} \\ &\stackrel{\text{Eq. 1}}{>} \frac{1}{\lambda} \sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i| - \left( 2 - \frac{1}{\lambda} \right) \cdot \frac{2\lambda}{T - 2\lambda} \cdot \sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i| \end{aligned}$$



$$\begin{aligned}
\Rightarrow \widehat{S}_{i_g} - \widehat{S}_{I_{sign}(\ell)} &= \frac{1}{\lambda} \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i| - \left( \frac{2\lambda-1}{\lambda} \right) \cdot \frac{2\lambda}{8\lambda^2-4\lambda} \cdot \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i| \\
&= \left( \frac{1}{\lambda} - \frac{2}{4\lambda} \right) \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i| \\
&= \frac{1}{2\lambda} \cdot \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i|
\end{aligned}$$

and

$$\begin{aligned}
\widehat{S}_{i_g} - \widehat{S}_{I_{sign}(\ell)} &\stackrel{\text{Lem. 3}}{\leq} 2 \cdot S_{i_g} - \frac{1}{\lambda} \cdot S_{I_{sign}(\ell)} = 2 \cdot \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i| + \left( 2 - \frac{1}{\lambda} \right) \cdot S_{I_{sign}(\ell)} \\
&\stackrel{\text{Eq. 1}}{\leq} 2 \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i| + \left( 2 - \frac{1}{\lambda} \right) \cdot \frac{2\lambda}{8\lambda^2-4\lambda} \cdot \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i| \\
&= \left( 2 + \frac{2}{4\lambda} \right) \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i| \\
&\leq \frac{5}{2} \cdot \sum_{i=i_g}^{I_{sign}(\ell)-1} |M_i|,
\end{aligned}$$

where we used  $\lambda \geq 1$  in the last inequality. Since  $|M_t| = S_t$  and  $\widehat{R}_t = \widehat{S}_t$ , the last statement follows directly from Lemma 3.  $\square$

From Lemma 2.1 we know that  $I_{sign} \subseteq I_{good}$  which together with the last Lemma 4 implies that  $\widehat{R}_{I_{sign}(\ell)}$  is a good estimator for  $D_\ell$ .

**Corollary 1.** *For  $\ell \in \{1, \dots, |I_{sign}|\}$ ,  $\frac{1}{2\lambda} \cdot D_\ell \leq \widehat{R}_{J(\ell)} \leq \frac{5}{2} \cdot D_\ell$ . Furthermore, if  $c > 5\lambda$  then the values of the  $D_\ell$  are exponentially increasing:*

$$D_1 \leq \frac{5\lambda}{c} D_2 \leq \dots \leq \left( \frac{5\lambda}{c} \right)^{|I_{sign}|-1} D_{|I_{sign}|-1}.$$

*Proof.* Recall that for  $\ell \in \{2, \dots, |I_{sign}|\}$  we defined  $D_\ell = \sum_{i=I_{sign}(\ell)}^{I_{sign}(\ell-1)+1} |M_i|$ . For  $\ell = 1$  the value of  $\widehat{R}_{I_{sign}(1)} = \widehat{S}_t$  is a good estimator for the size of the matching  $M_t$  (which is equal to  $D_1$ ) due to Lemma 3. Since for  $\ell \in \{2, \dots, |I_{sign}|\}$  it is  $\widehat{R}_{I_{sign}(\ell)} = S_{I_{sign}(\ell)} - S_{I_{sign}(\ell-1)}$  and  $I_{sign} \subseteq I_{good}$ , the first statement is a direct implication of Lemma 4 by setting  $i_g = I_{sign}(\ell)$ .

For three adjoining significant ranks  $I_{sign}(\ell+1)$ ,  $I_{sign}(\ell)$ ,  $I_{sign}(\ell-1)$  with  $\ell \in \{2, \dots, |I_{sign}|-1\}$ , we have

$$\begin{aligned}
\frac{1}{2\lambda} \cdot D_\ell &= \frac{1}{2\lambda} \sum_{i=I_{sign}(\ell)}^{I_{sign}(\ell-1)-1} |M_i| \stackrel{\text{Lem. 4}}{\leq} \widehat{S}_{I_{sign}(\ell)} - \widehat{S}_{I_{sign}(\ell-1)} = \widehat{R}_{I_{sign}(\ell)} \\
&\stackrel{\text{Lem. 2 (4)}}{\leq} \frac{1}{c} \cdot \widehat{R}_{I_{sign}(\ell+1)} = \frac{1}{c} \cdot \left( \widehat{S}_{I_{sign}(\ell+1)} - \widehat{S}_{I_{sign}(\ell)} \right) \\
&\stackrel{\text{Lem. 4}}{\leq} \frac{5}{2c} \sum_{i=I_{sign}(\ell+1)}^{I_{sign}(\ell)-1} |M_i| = \frac{5}{2c} \cdot D_{\ell+1}.
\end{aligned}$$

Since  $D_1 = |M_t|$  and  $\widehat{R_{I_{\text{sign}}(1)}} = \widehat{R_t} = \widehat{S_t}$ , Lemma 4 also implies that

$$\frac{1}{2\lambda} \cdot D_1 \leq \frac{1}{\lambda} \cdot D_1 \leq \widehat{R_t} \leq \frac{1}{c} \cdot \widehat{R_{I_{\text{sign}}(2)}} \leq \frac{5}{2c} \sum_{i=I_{\text{sign}}(2)}^{I_{\text{sign}}(1)-1} |M_i| = \frac{5}{2c} \cdot D_2.$$

Thus, for  $c > 5\lambda$  the values of the  $D_\ell$  are exponentially increasing:

$$D_1 \leq \frac{5\lambda}{c} D_2 \leq \dots \leq \left(\frac{5\lambda}{c}\right)^{|I_{\text{sign}}|-1} D_{|I_{\text{sign}}|-1}.$$

□

## (2) The Charging Argument

We show that the sum of the matching sizes between two significant ranks  $I_{\text{sign}}(\ell+1)$  and  $I_{\text{sign}}(\ell)$  is bounded by  $O(\lambda \cdot T \cdot D_\ell) = O\left(\lambda \cdot T \cdot \sum_{i=I_{\text{sign}}(\ell)}^{I_{\text{sign}}(\ell-1)+1} |M_i|\right)$ .

**Lemma 5.** *Setting  $c = \frac{2}{5} \cdot T + 5\lambda$  in Algorithm 2. Then for  $\ell \in \{1, \dots, |I_{\text{sign}}|-1\}$ ,*

$$\sum_{i=I_{\text{sign}}(\ell+1)+1}^{I_{\text{sign}}(\ell)-1} |M_i| \leq (2\lambda \cdot T + 25\lambda^2) \cdot D_{|I_{\text{sign}}|} \text{ if } 0 \notin I_{\text{sign}}.$$

*and*

$$\sum_{i=0}^{I_{\text{sign}}(|I_{\text{sign}}|)-1} |M_i| \leq (2\lambda \cdot T + 25\lambda^2) \cdot D_{|I_{\text{sign}}|} \text{ if } 0 \notin I_{\text{sign}}.$$

*Proof.* For the proof of the first inequality, let  $i_g \in I_{\text{good}}$  be minimal such that  $I_{\text{sign}}(\ell+1) < i_g < I_{\text{sign}}(\ell)$  for  $\ell \in \{1, \dots, |I_{\text{sign}}|-1\}$ . If such a good rank does not exist, set  $i_g = -1$ . We distinguish between two cases. Note that  $c = \frac{5}{2} \cdot T + 5\lambda > 5\lambda$ .

**Case 1:**  $i_g = I_{\text{sign}}(\ell+1) + 1$ . For the sake of simplicity, we abuse the notation and set  $\widehat{S_{I_{\text{sign}}(0)}} = 0$  such that  $\widehat{R_{I_{\text{sign}}(\ell)}} = \widehat{S_{I_{\text{sign}}(\ell)}} - \widehat{S_{I_{\text{sign}}(\ell-1)}}$  also holds for  $\ell = 1$ . Using Lemma 4 we have

$$\begin{aligned} \sum_{i=I_{\text{sign}}(\ell+1)+1}^{I_{\text{sign}}(\ell)-1} |M_i| &= \sum_{i=i_g}^{I_{\text{sign}}(\ell)-1} |M_i| \stackrel{\text{Lem. 4}}{\leq} 2\lambda \cdot \left(\widehat{S_{i_g}} - \widehat{S_{I_{\text{sign}}(\ell)}}\right) \\ &\stackrel{i_g \notin I_{\text{sign}}}{\leq} 2\lambda c \cdot \widehat{R_{I_{\text{sign}}(\ell)}} = 2\lambda \cdot c \cdot \left(\widehat{S_{I_{\text{sign}}(\ell)}} - \widehat{S_{I_{\text{sign}}(\ell-1)}}\right) \\ &\stackrel{\text{Lem. 4}}{\leq} 5\lambda \cdot c \cdot \sum_{i=I_{\text{sign}}(\ell)}^{I_{\text{sign}}(\ell-1)-1} |M_i| = 5 \cdot c \cdot D_\ell \end{aligned} \quad (2)$$

**Case 2:**  $i_g \neq I_{\text{sign}}(\ell+1) + 1$ . In this case  $\widehat{S_{I_{\text{sign}}(\ell+1)+1}} \leq T \cdot \widehat{S_{I_{\text{sign}}(\ell)}}$ . Thus

$$\begin{aligned} \sum_{i=I_{\text{sign}}(\ell+1)+1}^{I_{\text{sign}}(\ell)-1} |M_i| &\leq \widehat{S_{I_{\text{sign}}(\ell+1)+1}} \stackrel{\text{Lem. 3}}{\leq} \lambda \cdot \widehat{S_{I_{\text{sign}}(\ell+1)+1}} \\ &\leq \lambda \cdot T \cdot \widehat{S_{I_{\text{sign}}(\ell)}} \stackrel{\text{Lem. 3}}{\leq} 2\lambda \cdot T \cdot S_{I_{\text{sign}}(\ell)} = 2\lambda \cdot T \cdot \sum_{i=1}^{\ell} D_i \\ &\stackrel{\text{Cor. 1}}{\leq} 2\lambda \cdot T \cdot D_\ell \cdot \sum_{i=1}^{\ell} \left(\frac{5\lambda}{c}\right)^i \leq 2\lambda \cdot T \cdot D_\ell \cdot \frac{1}{1 - \frac{5\lambda}{c}} \end{aligned} \quad (3)$$

Combining the inequalities 2 and 3, we have  $\sum_{i=I_{\text{sign}}(\ell+1)+1}^{I_{\text{sign}}(\ell)-1} |M_i| \leq \max \left\{ 5\lambda \cdot c, \frac{2\lambda \cdot T}{1 - \frac{5\lambda}{c}} \right\} \cdot D_\ell$  which simplifies to

$$\sum_{i=I_{\text{sign}}(\ell+1)+1}^{I_{\text{sign}}(\ell)-1} |M_i| \leq (2\lambda \cdot T + 25\lambda^2) \cdot D_\ell \quad \text{for } \ell \in \{1, \dots, |I_{\text{sign}}| - 1\}.$$

If  $0 \notin I_{\text{sign}}$  we can do the same arguments to bound  $\sum_{i=0}^{I_{\text{sign}}(|I_{\text{sign}}|)-1} |M_i|$  by  $(2\lambda \cdot T + 25\lambda^2) \cdot D_{|I_{\text{sign}}|}$ . Let  $i_g \in I_{\text{good}}$  be minimal such that  $0 \leq i_g < I_{\text{sign}}(|I_{\text{sign}}|)$ . Again, we distinguish between two cases.

**Case 1:**  $i_g = 0$ . Using Lemma 4 we have

$$\begin{aligned} \frac{1}{2\lambda} \cdot \sum_{i=0}^{I_{\text{sign}}(|I_{\text{sign}}|)-1} |M_i| &\stackrel{\text{Lem. 4}}{\leq} \widehat{S_0} - \widehat{S_{I_{\text{sign}}(|I_{\text{sign}}|)}} \\ &\stackrel{0 \notin I_{\text{sign}}}{\leq} c \cdot R_{I_{\text{sign}}(|I_{\text{sign}}|)} = c \cdot \left( \widehat{S_{I_{\text{sign}}(|I_{\text{sign}}|)}} - \widehat{S_{I_{\text{sign}}(|I_{\text{sign}}|-1)}} \right) \\ &\stackrel{\text{Lem. 4}}{\leq} \frac{5}{2} \cdot c \cdot \sum_{i=I_{\text{sign}}(|I_{\text{sign}}|)}^{I_{\text{sign}}(|I_{\text{sign}}|-1)-1} |M_i| = \frac{5}{2} \cdot c \cdot D_{|I_{\text{sign}}|} \\ \Leftrightarrow \sum_{i=0}^{I_{\text{sign}}(|I_{\text{sign}}|)-1} |M_i| &\leq 5\lambda \cdot c \cdot D_{|I_{\text{sign}}|} \end{aligned}$$

**Case 2:**  $i_g \neq 0$ . In this case  $\widehat{S_0} \leq T \cdot \widehat{S_{I_{\text{sign}}(|I_{\text{sign}}|)}}$ . Thus

$$\begin{aligned} \frac{1}{\lambda} \cdot \sum_{i=0}^{I_{\text{sign}}(|I_{\text{sign}}|)-1} |M_i| &\leq \frac{1}{\lambda} \cdot S_0 \stackrel{\text{Lem. 3}}{\leq} \widehat{S_0} \leq T \cdot \widehat{S_{I_{\text{sign}}(|I_{\text{sign}}|)}} \\ &\stackrel{\text{Lem. 3}}{\leq} 2 \cdot T \cdot S_{I_{\text{sign}}(|I_{\text{sign}}|)} = 2 \cdot T \cdot \sum_{i=1}^{|I_{\text{sign}}|} D_i \\ &\stackrel{\text{Cor. 1}}{\leq} 2 \cdot T \cdot D_{|I_{\text{sign}}|} \cdot \sum_{i=1}^{|I_{\text{sign}}|} \left( \frac{5\lambda}{c} \right)^i \leq 2 \cdot T \cdot D_{|I_{\text{sign}}|} \cdot \frac{1}{1 - \frac{5\lambda}{c}} \\ \Leftrightarrow \sum_{i=0}^{I_{\text{sign}}(|I_{\text{sign}}|)-1} |M_i| &\leq \frac{2\lambda \cdot T}{1 - \frac{5\lambda}{c}} \cdot D_{|I_{\text{sign}}|}. \end{aligned}$$

Now, with the same  $c = \frac{2}{5} \cdot T + 5\lambda$  as before we have

$$\sum_{i=0}^{I_{\text{sign}}(|I_{\text{sign}}|)-1} |M_i| \leq (2\lambda \cdot T + 25\lambda^2) \cdot D_{|I_{\text{sign}}|}.$$

□

We use Lemma 5 to show that  $w(M)$  is bounded in terms of  $\sum_{\ell=1}^{|I_{\text{sign}}|} r_{I_{\text{sign}}(\ell)} \cdot D_\ell$ :

$$\sum_{i=0}^t r_i \cdot |M_i| \geq \sum_{\ell=1}^{|I_{\text{sign}}|} r_{I_{\text{sign}}(\ell)} \cdot D_\ell \quad (4)$$

$$\sum_{i=0}^t r_i \cdot |M_i| \leq (1 + 2\lambda \cdot T + 25\lambda^2) \cdot \sum_{\ell=1}^{|I_{\text{sign}}|} r_{I_{\text{sign}}(\ell)} \cdot D_\ell. \quad (5)$$

### Putting Everything Together

Using Corollary 1 we have  $\frac{1}{2\lambda} \cdot D_\ell \leq \widehat{R_{I_{\text{sign}}(\ell)}} \leq \frac{5}{2} \cdot D_\ell$  for all  $\ell \in \{1, \dots, |I_{\text{sign}}|\}$  which with (4) and (5) gives  $\frac{1}{2\lambda \cdot (1 + 2\lambda \cdot T + 25\lambda^2)} \cdot w(M) \leq \sum_{\ell=1}^{|I_{\text{sign}}|} r_{I_{\text{sign}}(\ell)} \cdot \widehat{R_{I_{\text{sign}}(\ell)}} \leq \frac{5}{2} \cdot w(M)$ . Recall that we set  $T = 8\lambda^2 - 2\lambda$ . Now, folding in the factor of  $\frac{1}{8}$  from the partitioning and rescaling the estimator gives an  $O(\lambda^4)$ -estimation on the weight of an optimal weighted matching.  $\square$

## 2.1 Applications

Since every edge insertion and deletion supplies the edge weight, it is straightforward to determine the rank for each edge upon every update. Using the following results for unweighted matching, we can obtain estimates with similar approximation guarantee and space bounds for weighted matching.

### Random Order Streams

For an arbitrary graph whose edges are streamed in random order, Kapralov, Khanna and Sudan [23] gave an algorithm with polylog  $n$  approximation guarantee using polylog  $n$  space with failure probability  $\delta = 1/\text{polylog } n$ . Since this probability takes the randomness of the input permutation into account, we cannot easily amplify it, though for  $\log W \leq \delta$ , the extension to weighted matching still succeeds with at least constant probability.

### Adversarial Streams

The arboricity of a graph  $G$  is defined as  $\max_{U \subseteq V} \left\lceil \frac{|E(U)|}{|U|-1} \right\rceil$ . Examples of graphs with constant arboricity include planar graphs and graphs with constant degree. For graphs of bounded arboricity  $\nu$ , Esfandiari et al. [15] gave an algorithm with an  $O(\nu)$  approximation guarantee using  $\tilde{O}(\nu \cdot n^{2/3})$  space.

### Dynamic Streams

We give two estimation algorithms for the size of a maximum matching. First, we see that it is easy to estimate the matching size in trees. Second, we extend the result from [15] where the matching size of so called bounded arboricity graphs in insertion-only streams is estimated to dynamic graph streams.

**Matching Size of Trees** Let  $T = (V, E)$  be a tree with at least 3 nodes and let  $h_T$  be the number of internal nodes, i.e. nodes with degree greater than 1. We know that the size of a maximum matching is between  $h_T/2$  and  $h_T$ . Therefore, it suffices to estimate the number of internal nodes of a tree to approximate the maximum matching within  $2 + \varepsilon$  factor which was also observed in [15]. In order to estimate the matching size, we maintain an  $\ell_0$ -Estimator for the degree vector  $d \in \mathbb{R}^N$  such that  $d_v = \deg(v) - 1$  holds at the end of the stream and with it  $\ell_0(d) = h_T$ . In

other words, we initialize the vector by adding  $-1$  to each entry and update the two corresponding entries when we get an edge deletion or insertion. Since the number of edges in a tree is  $N - 1$ , the preprocessing time can be amortized during the stream. Using Theorem 10 from Kane et al. [21], we can maintain the  $\ell_0$ -Estimator for  $d$  in  $O(\varepsilon^{-2} \log^2 N)$  space.

**Theorem 3.** *Let  $T = (V, E)$  be a tree with at least 3 nodes and let  $\varepsilon \in (0, 1)$ . Then there is an algorithm that estimates the size of a maximum matching in  $T$  within a  $(2 + \varepsilon)$ -factor in the dynamic streaming model using 1-pass over the data and  $O(\varepsilon^{-2} \log^2 N)$  space.*

As in [15] this algorithm can be extended to forests with no isolated node.

**Matching Size in Graphs with Bounded Arboricity** The algorithm is based on the results from [15]. Since we need parametrized versions of their results, we summarize and rephrase the ideas and proofs in this section. Let  $G = (V, E)$  be a graph. The arboricity  $a(G)$  of  $G$  is a kind of density measure: The number of edges in every induced subgraph of size  $s$  in  $G$  is bounded by  $s \cdot a(G)$ . Formally, the arboricity  $a(G)$  of  $G$  is defined by  $a(G) = \max_{U \subseteq V} \left\lceil \frac{|E(U)|}{|U|-1} \right\rceil$ . If  $\mu_G$  is an upper bound on the average degree of every induced subgraph of  $G$  then  $\mu_G \leq 2 \cdot a(G)$ .

**Definition 2** ([15]). A node  $v \in V$  is *light* if  $\deg(v) \leq C$  with  $C = \lceil \mu_G \rceil + 3$ . Otherwise,  $v$  is *heavy*. An edge is *shallow* if and only if both of its endpoints are light. We denote by  $h_G$  the number of heavy nodes in  $G$  and by  $s_G$  the number of shallow edges in  $G$ , respectively.

Using the results from Czygrinow, Hanchowiak, and Szymanska [12] (and  $C = 20a(G)/\varepsilon^2$ ) it is possible to get a  $O(a(G))$  approximation for the size of a maximum matching by just estimating  $h_G$  and  $s_G$ . Esfandiari et al. [15] improved the approximation factor to roughly  $5 \cdot a(G)$ .

**Lemma 6** ([15]). *Let  $G = (V, E)$  be a graph with maximum matching  $M^*$ . Then we have  $\frac{\max\{h_G, s_G\}}{\eta} \leq |M^*| \leq h_G + s_G$  where  $\eta = 1.25C + 0.75$  where  $C$  is at most  $\lceil 2a(G) + 3 \rceil$ .*

Estimating  $h_G$  and  $s_G$  is possible by random sampling: For heavy nodes, we randomly draw a large enough set of nodes and count the heavy nodes by maintaining their degree. Rescaling the counter gives a sufficiently good estimate, provide  $h_G$  is large enough. For  $s_G$  we randomly draw nodes and maintain the induced subgraph. For each node contained in the subgraph it is straightforward to maintain the degree and thereby to decide whether or not a given edge from the subgraph is shallow. Then we can rescale the counted number of shallow edges which gives us an estimation on  $s_G$  if  $s_G$  is large enough. Dealing with small values of  $s_G$  and  $h_G$ , Esfandiari et al. additionally maintain a small maximal matching of size at most  $n^\alpha$  with  $\alpha < 1$ . If the maintained matching exceeds this value then we know that either  $s_G$  or  $h_G$  is greater than  $n^\alpha/2$  by Lemma 6 and the estimation of the parameters  $h_G$  and  $s_G$  will be sufficiently accurate. The main tool to extend this algorithm to dynamic graph streams is to estimate the size of a small matching by means of the Tutte matrix. But first, we restate the following three lemmas from [15] for arbitrary parameters and extend them to dynamic streams.

**Lemma 7.** *Let  $T$  be an integer and  $\varepsilon \leq 1/\sqrt{3}$ . Then there exists a 1-pass algorithm for dynamic streams that outputs a value  $h$  which is a  $(1 \pm \varepsilon)$  estimation of  $h_G$  if  $h_G \geq T$  and which is smaller than  $3T$  otherwise. The algorithm needs  $O\left(\frac{\log^2 n}{\varepsilon^2} \cdot \frac{n}{T}\right)$  space and succeeds with high probability.*

*Proof.* The probability of sampling a heavy node is  $\frac{h_G}{n}$ . Hence, sampling a set of nodes  $S$  gives us  $|S| \cdot \frac{h_G}{n}$  heavy nodes on expectation. Set  $|S| = \frac{3 \log n}{\varepsilon^2} \cdot \frac{n}{T}$ . For each node  $v \in S$  we maintain its degree using  $O(\log N)$  space. We define the indicator variable  $X_v$  with  $v \in S$  which is 1 if  $v$  is heavy and

0 otherwise. Then our estimator for  $h_G$  is  $\hat{h} = \frac{n}{|S|} \sum X_v$  which is equal to  $h_G$  in expectation. First, assume  $h_G \geq T$ . Then using the Chernoff bound we have

$$\begin{aligned} \mathbb{P} \left[ \hat{h} \geq (1 + \varepsilon) \cdot \mathbb{E} [\hat{h}] \right] &= \mathbb{P} \left[ \sum_{v \in S} X_v \geq (1 + \varepsilon) \cdot \mathbb{E} \left[ \sum_{v \in S} X_v \right] \right] \\ &\leq \exp \left( -\frac{3 \log n}{\varepsilon^2} \frac{n}{T} \cdot \frac{h_G}{n} \cdot \frac{\varepsilon^2}{3} \right) \leq \frac{1}{n}. \end{aligned}$$

The same bound also holds for  $\mathbb{P} \left[ \hat{h} \leq (1 - \varepsilon) \cdot \mathbb{E} [\hat{h}] \right]$ . If  $h_G < T$ , then again using the Chernoff bound gives us

$$\begin{aligned} &\mathbb{P} \left[ \frac{n}{|S|} \cdot \left( \sum_{v \in S} X_v \right) \geq 3T \right] \\ &= \mathbb{P} \left[ \sum_{v \in S} X_v \geq \frac{3T \cdot |S| \cdot h_G}{n \cdot h_G} \right] \\ &= \mathbb{P} \left[ \sum_{v \in S} X_v > \left( 1 + \frac{3T}{h_G} - 1 \right) \cdot \mathbb{E} \left[ \sum_{v \in S} X_v \right] \right] \\ &\leq \exp \left( -\frac{3 \log n}{\varepsilon^2} \frac{n}{T} \cdot \frac{h_G}{n} \cdot \frac{\frac{3T}{h_G} - 1}{2} \right) \\ &\leq \exp \left( -\frac{3 \log n}{\varepsilon^2} \frac{n}{T} \cdot \frac{h_G}{n} \cdot \frac{\frac{2T}{h_G}}{2} \right) \leq \frac{1}{n}, \end{aligned}$$

where the last inequality follows from  $\varepsilon \leq \frac{1}{\sqrt{3}}$ .  $\square$

**Lemma 8.** *Let  $T$  be an integer and  $\varepsilon \leq 1/\sqrt{3}$ . Then there exists a 2-pass algorithm for dynamic streams that outputs a value  $\hat{s}$  which is a  $(1 \pm \varepsilon)$  estimation of  $s_G$  if  $s_G \geq T$  and which is smaller than  $3T$  if  $s_G < T$ . The algorithm uses  $O\left(\frac{a(G) \cdot n \log^4 n}{\varepsilon^2 T}\right)$  space and succeeds with high probability.*

*Proof.* In the first pass, we sample  $\frac{3 \log n}{\varepsilon^2} \frac{a(G) \cdot n}{T}$  edges uniformly at random using  $\ell_0$  samplers, each of which cost at most  $O(\log^3 n)$  space [20]. For each node of a sampled edge, we maintain its degree in the second pass to decide whether a given edge is shallow or not. Hereafter, we reapply the analysis of Lemma 7: Let  $S = (e_1, \dots, e_{|S|})$  be the sequence of sampled edges in the first pass and let  $X_i$  be the indicator variable which is 1 if and only if  $e_i$  is shallow. The probability of sampling a shallow edge is  $\frac{s_G}{|E|}$  which implies that  $\mathbb{E}[\sum X_i] = |S| \cdot \frac{s_G}{|E|} \geq |S| \cdot \frac{s_G}{a(G) \cdot n}$ . Now, let  $\hat{s} = \frac{|E|}{|S|} \sum X_i$  be our estimator. We know that  $\mathbb{E}[\hat{s}] = s_G$ . If  $s_G \geq T$  then by Chernoff we have

$$\begin{aligned} \mathbb{P} [\hat{s} \geq (1 + \varepsilon) \cdot \mathbb{E} [\hat{s}]] &= \mathbb{P} \left[ \sum X_i \geq (1 + \varepsilon) \cdot \mathbb{E} \left[ \sum X_i \right] \right] \\ &\leq \exp \left( -\frac{3 \log n}{\varepsilon^2} \frac{a(G) \cdot n}{T} \cdot \frac{s_G}{a(G) \cdot n} \cdot \frac{\varepsilon^2}{3} \right) \leq \frac{1}{n}. \end{aligned}$$

The same bound also holds for  $\mathbb{P} [\hat{s} \leq (1 - \varepsilon) \cdot \mathbb{E} [\hat{s}]]$ . If  $s_G < T$ , then again using the Chernoff

bound gives us

$$\begin{aligned}
& \mathbb{P} \left[ \frac{|E|}{|S|} \cdot \left( \sum X_i \right) \geq 3T \right] \\
&= \mathbb{P} \left[ \sum X_i \geq \frac{3T \cdot |S| \cdot s_G}{|E| \cdot s_G} \right] \\
&= \mathbb{P} \left[ \sum X_i > \left( 1 + \frac{3T}{s_G} - 1 \right) \cdot \mathbb{E} \left[ \sum X_i \right] \right] \\
&\leq \exp \left( -\frac{3 \log n}{\varepsilon^2} \frac{a(G) \cdot n}{T} \cdot \frac{s_G}{a(G) \cdot n} \cdot \frac{\frac{3T}{s_G} - 1}{2} \right) \\
&\leq \exp \left( -\frac{3 \log n}{\varepsilon^2} \frac{a(G) \cdot n}{T} \cdot \frac{s_G}{a(G) \cdot n} \cdot \frac{\frac{2T}{s_G}}{2} \right) \leq \frac{1}{n},
\end{aligned}$$

where the last inequality follows from  $\varepsilon \leq \frac{1}{\sqrt{3}}$ .  $\square$

**Lemma 9.** *Let  $\varepsilon > 0$  and  $T > (16C/\varepsilon)^2$  be an integer. Then there exists a 1-pass algorithm for dynamic streams that outputs a value  $\hat{s}$  which is a  $(1 \pm \varepsilon)$  estimation of  $s_G$  if  $s_G \geq T$  and which is smaller than  $3T$  if  $s_G < T$ . The algorithm uses  $\tilde{O} \left( \frac{a(G) \cdot n}{\varepsilon \sqrt{T}} \right)$  space and succeeds with constant probability.*

*Proof.* Let  $S$  be a set of  $\frac{4n}{\varepsilon \sqrt{T}}$  randomly chosen nodes. We maintain the entire subgraph induced by  $S$  and the degree of each node in  $S$ . Note that the number of edges in this subgraph at the end of the stream is at most  $a(G) \cdot |S|$ . Since we have edge deletions this number may be exceeded at some point during the stream. Thus, we cannot explicitly store the subgraph but we can recover all entries using an  $a(G) \cdot |S|$ -sparse recovery sketch using  $\tilde{O}(a(G) \cdot |S|)$  space (see Barkay et al [7]). Let  $e_1, \dots, e_{s_G}$  be the shallow edges in  $G$ . Define  $X_i = 1$  if  $e_i \in E(S)$  and 0 otherwise.  $X_i$  is Bernoulli distributed where the probability of both nodes being included in the subgraph follows from the hypergeometric distribution with population  $n$ , 2 successes in the population, sample size  $|S|$  and 2 successes in the sample:

$$p = \frac{\binom{2}{2} \binom{n-2}{|S|-2}}{\binom{n}{|S|}} = \frac{|S| \cdot (|S| - 1)}{n \cdot (n - 1)} \geq \frac{|S|^2}{2n^2} = \frac{8}{\varepsilon^2 T}.$$

Hence  $X_i$  is Bernoulli distributed, we have  $\text{Var}[X_i] = p \cdot (1 - p) \leq p$ . We know that  $\text{Var}[\sum X_i] = \sum \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}[X_i, X_j]$ . For the covariance between two variables  $X_i$  and  $X_j$  we have two cases: If  $e_i$  and  $e_j$  do not share a node, then  $X_i$  and  $X_j$  cannot be positively correlated, i.e.  $\text{Cov}[X_i, X_j] > 0$ . To be more precise, we observe that by definition  $\text{Cov}[X_i, X_j]$  is equal to  $\mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \cdot \mathbb{E}[X_j]$  which is equal to  $\mathbb{P}[X_i = X_j = 1] - p^2$ . The probability  $\mathbb{P}[X_i = X_j = 1]$  is equal to the probability of drawing exactly 4 fixed nodes from  $V$  with a sample of size  $|S|$  which is

$$\frac{\binom{4}{4} \binom{n-4}{|S|-4}}{\binom{n}{|S|}} = \frac{|S| \cdot (|S| - 1) \cdot (|S| - 2) \cdot (|S| - 3)}{n \cdot (n - 1) \cdot (n - 2) \cdot (n - 3)}.$$

Since  $\frac{a+c}{b+c} \geq \frac{a}{b}$  for  $a \leq b$  and  $c \geq 0$ , this probability is at most  $p^2$  which means that the covariance is at most 0. If  $e_i$  and  $e_j$  share a node, we have

$$\begin{aligned}
\text{Cov}[X_i, X_j] &\leq \mathbb{P}[X_i = X_j = 1] \\
&= \frac{\binom{3}{3} \binom{n-3}{|S|-3}}{\binom{n}{|S|}} = \frac{|S| \cdot (|S| - 1) \cdot (|S| - 2)}{n \cdot (n - 1) \cdot (n - 2)} \leq p^{3/2}.
\end{aligned}$$

By definition each node incident to a shallow edge has at most  $C$  neighbors and therefore, we have at most  $2C$  edges that share a node with a given shallow edge. In total, we can bound the variance of  $X$

$$\begin{aligned} \text{Var}[X] &= \sum_i \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}[X_i, X_j] \\ &\leq p \cdot s_G + \sum_{\substack{e_i \neq e_j, \\ e_i, e_j \text{ share a node}}} \text{Cov}[X_i, X_j] \leq p \cdot s_G + 2C \cdot s_G \cdot p^{3/2} \leq 2p \cdot s_G \end{aligned}$$

where the last inequality follows from  $\sqrt{p} \leq \frac{|S|}{N/2} = \frac{8}{\varepsilon\sqrt{T}}$  and  $T \geq (16C/\varepsilon)^2$ . Using Chebyshev's inequality we have for  $s_G \geq T$

$$\begin{aligned} \mathbb{P}\left[\left|\frac{1}{p} \cdot X - \frac{1}{p} \mathbb{E}[X]\right| > \epsilon \cdot \frac{1}{p} \mathbb{E}[X]\right] &= \mathbb{P}[|X - \mathbb{E}[X]| > \epsilon \cdot \mathbb{E}[X]] \\ &\leq \frac{\text{Var}[X]}{\epsilon^2 \mathbb{E}[X]^2} \leq \frac{2p \cdot s_G}{\epsilon^2 p^2 \cdot s_G^2} \leq \frac{2}{\epsilon^2 T p} \\ &\leq \frac{2\varepsilon^2 T}{8\epsilon^2 T} = \frac{1}{4}. \end{aligned}$$

If  $s_G < T$ , we have  $\mathbb{E}[X] = p \cdot s_G < pT$ . Thus, it is

$$\begin{aligned} \mathbb{P}\left[\frac{1}{p} \cdot X \geq 3T\right] &= \mathbb{P}[X - \mathbb{E}[X] \geq 3Tp - \mathbb{E}[X]] \\ &\leq \mathbb{P}[|X - \mathbb{E}[X]| \geq 2Tp] \\ &\leq \frac{\text{Var}[X]}{4T^2 p^2} \leq \frac{2p \cdot s_G}{4T^2 p^2} \leq \frac{2}{4Tp} \leq \frac{2\varepsilon^2 T}{16T} = \frac{\varepsilon^2}{16} \leq \frac{1}{16}. \end{aligned}$$

□

---

### Algorithm 3 Unweighted Matching Approximation

---

**Require:**  $G = (V, E)$  with  $a(G) \leq \alpha$  and  $\varepsilon \in (0, 1/\sqrt{3})$

**Ensure:** Estimator on the size of a maximum matching

Set  $T = n^{2/5}$  for a single pass and  $T = n^{1/3}$  for two passes and  $\eta = 2.5\lceil 2 \cdot \alpha + 3 \rceil + 5.75$ .

Let  $\hat{h}$  and  $\hat{s}$  be the estimators from Lemma 7 and Lemma 9

**for**  $i = 0, \dots, \log 3T/(1 - \varepsilon)$  **do**

Solve rank decision with parameter  $k = 2^i$  on the Tutte-Matrix  $T(G)$  with randomly chosen indeterminates

**if**  $\text{rank}(T(G)) < 3T/(1 - \varepsilon)$  **then**

Output the maximal  $2^{i+1}$  for the maximal  $i \in \{0, \dots, \log 3T/(1 - \varepsilon)\}$  with  $\text{rank}(T(G)) \geq 2^i$

**else**

Output  $\frac{\max\{\hat{h}, \hat{s}\}}{(1 + \varepsilon)\eta}$ .

---

Algorithm 3 shows the idea of the estimation of the unweighted maximum matching size in bounded arboricity graphs using the previous results and the relation between the rank of the Tutte matrix and the matching size.

**Theorem 4.** *Let  $G$  be a graph with  $a(G) \leq \alpha$  with  $n \geq (16\alpha/\varepsilon)^5$ . Let  $\varepsilon \in (0, 1/\sqrt{3})$ . Then there exists an algorithm estimating the size of the maximum matching in  $G$  within a  $\frac{2(1+\varepsilon)(5 \cdot a(G) + O(1))}{(1-\varepsilon)}$  factor in the dynamic streaming model using*



- a single pass over the data and  $\tilde{O}(\frac{\alpha \cdot n^{4/5}}{\varepsilon^2})$  space or
- 2 passes over the data and  $\tilde{O}(\alpha \cdot n^{2/3})$  space.

*Proof.* For the sake of simplicity we assume that  $3T/(1 - \varepsilon)$  is a power of two. We know that we can decide the rank decision problem with parameter  $k$  in a dynamic stream with one pass using  $O(k^2 \log n)$  space by Theorem 5.1 of Clarkson and Woodruff [9]. Thus, invoking this algorithm for  $k = 2^0, 2^1, \dots, 2^{\log 3T/(1-\varepsilon)}$  results in a space requirement of  $O(T^2 \cdot \log T \cdot \log n) = O(T^2 \log^2 N)$  for our choices of  $T$ . For the first part of the theorem, we estimate  $s_G$  and  $h_G$  in 1-pass by  $\hat{h}$  and  $\hat{s}$  using  $\tilde{O}(\frac{n}{\varepsilon^2 T})$  and  $\tilde{O}(\frac{\alpha \cdot n}{\varepsilon \sqrt{T}})$  space, see Lemma 7 and Lemma 9. Setting  $T = n^{2/5}$  gives us the desired space bound of  $\tilde{O}(\frac{\alpha \cdot n^{4/5}}{\varepsilon^2})$  (note that  $T > (16\alpha/\varepsilon)^2$  which is required for Lemma 9). For the second part of the theorem, we can improve the space requirements for the estimator  $\hat{h}$  and  $\hat{s}$  to  $\tilde{O}(\frac{a(G)n}{T})$  by using Lemma 7 and Lemma 8. Now, setting  $T = n^{1/3}$  gives the desired space bound.

Let  $OPT$  be the size of a maximum matching. First, we check whether  $OPT \geq 2 \cdot 3T/(1 - \varepsilon)$  by invoking the rank decision algorithm with parameter  $k = 3T/(1 - \varepsilon)$ . Since the rank of the matrix is equal to  $2OPT$ , this decides whether  $OPT \geq 2 \cdot 3T/(1 - \varepsilon)$ . If this is not true, we can give a 2-approximation on  $OPT$  by testing whether the rank of the Tutte matrix is in  $[2^i, 2^{i+1})$  for  $i = 0, \dots, \log(3T/(1 - \varepsilon)) - 1$ . If  $OPT \geq 2 \cdot 3T/(1 - \varepsilon)$  Lemma 6 implies that  $\max\{h_G, s_G\} \geq 3T/(1 - \varepsilon)$  since  $h_G + s_G \geq OPT$ . Assuming that we can approximate  $\max\{h_G, s_G\}$  then again by Lemma 6 we can estimate  $OPT$  since

$$\frac{\max\{h_G, s_G\}}{\eta} \leq OPT \leq h_G + s_G \leq 2 \max\{h_G, s_G\}.$$

W.l.o.g. let  $\hat{h} = \arg \max\{\hat{h}, \hat{s}\}$ . Now we have two cases:

1. If  $h_G = \arg \max\{h_G, s_G\} \geq T$  then by Lemma 7  $\hat{h}$  is a  $(1 \pm \varepsilon)$  estimation on  $h_G$ .
2. If  $s_G = \arg \max\{h_G, s_G\} \geq 3T/(1 - \varepsilon)$  we know by Lemma 9 that  $\hat{s} \geq 3T$  which implies that  $\hat{h} \geq \hat{s} \geq 3T$ . Thus by Lemma 7  $\hat{h}$  is a  $(1 \pm \varepsilon)$  estimation on  $h_G$ . This gives us

$$(1 - \varepsilon)s_G \leq \hat{s} \leq \hat{h} \leq (1 + \varepsilon)h_G \leq (1 + \varepsilon)s_G.$$

Therefore,  $\max\{\hat{h}, \hat{s}\}$  is a good estimator for  $\max\{h_G, s_G\}$ . For the estimator  $\frac{\max\{\hat{h}, \hat{s}\}}{(1 + \varepsilon)\eta}$  we have

$$\frac{(1 - \varepsilon)}{2(1 + \varepsilon)\eta} \cdot OPT \leq \frac{(1 - \varepsilon) \max\{h_G, s_G\}}{(1 + \varepsilon)\eta} \leq \frac{\max\{\hat{h}, \hat{s}\}}{(1 + \varepsilon)\eta} \leq \frac{(1 + \varepsilon) \max\{h_G, s_G\}}{(1 + \varepsilon)\eta} \leq OPT.$$

□

### 3 Lower Bound

Esfandiari et al. [15] showed a space lower bound of  $\Omega(\sqrt{n})$  for any estimation better than  $3/2$ . Their reduction (see below) uses the Boolean Hidden Matching Problem introduced by Bar-Yossef et al. [6], and further studied by Gavinsky et al. [17]. We will use the following generalization due to Verbin and Yu [33].

**Definition 3** (Boolean Hidden Hypermatching Problem [33]). In the *Boolean Hidden Hypermatching* Problem  $BHH_{t,n}$  Alice gets a vector  $x \in \{0, 1\}^n$  with  $n = 2kt$  and  $k \in \mathbb{N}$  and Bob gets a perfect  $t$ -hypermatching  $M$  on the  $n$  coordinates of  $x$ , i.e., each edge has exactly  $t$  coordinates, and a string

$w \in \{0, 1\}^{n/t}$ . We denote the vector of length  $n/t$  given by  $(\bigoplus_{1 \leq i \leq t} x_{M_{1,i}}, \dots, \bigoplus_{1 \leq i \leq t} x_{M_{n/t,i}})$  by  $Mx$  where  $(M_{1,1}, \dots, M_{1,t}), \dots, (M_{n/t,1}, \dots, M_{n/t,t})$  are the edges of  $M$ . The problem is to return 1 if  $Mx \oplus w = 1^{n/t}$  and 0 if  $Mx \oplus w = 0^{n/t}$ , otherwise the algorithm may answer arbitrarily.

Verbin and Yu [33] showed a lower bound of  $\Omega(n^{1-1/t})$  for the randomized one-way communication complexity for  $BHH_{t,n}$ . For our reduction we require  $w = 0^{n/t}$  and thus  $Mx = 1^{n/t}$  or  $Mx = 0^{n/t}$ . We denote this problem by  $BHH_{t,n}^0$ . We can show that this does not reduce the communication complexity.

**Lemma 10.** *The communication complexity of  $BHH_{t,4n}^0$  is lower bounded by the communication complexity of  $BHH_{t,n}$ .*

*Proof.* First, let assume that  $t$  is odd. Let  $x \in \{0, 1\}^n$  with  $n = 2kt$  for some  $k \in \mathbb{N}$  and  $M$  be a perfect  $t$ -hypermatching on the  $n$  coordinates of  $x$  and  $w \in \{0, 1\}^{n/t}$ . We define  $x' = [x^T x^T \bar{x}^T \bar{x}^T]^T$  to be the concatenation of two identical copies of  $x$  and two identical copies of the vector resulting from the bitwise negation of  $x$ . W.l.o.g. let  $\{x_1, \dots, x_t\} \in M$  be the  $l$ -th hyperedge of  $M$ . Then we add the following four hyperedges to  $M'$ :

- $\{x_1, \bar{x}_2, \dots, \bar{x}_t\}, \{\bar{x}_1, x_2, \bar{x}_3, \dots, \bar{x}_t\}, \{\bar{x}_1, \bar{x}_2, x_3, \dots, x_t\}$ , and  $\{x_1, \dots, x_t\}$  if  $w_l = 0$ ,
- $\{\bar{x}_1, x_2, \dots, x_t\}, \{x_1, \bar{x}_2, \dots, x_t\}, \{x_1, x_2, \bar{x}_3, \dots, \bar{x}_t\}$ , and  $\{\bar{x}_1, \dots, \bar{x}_t\}$  if  $w_l = 1$ .

The important observation here is that we flip even number of bits in the case  $w_l = 0$  and an odd number of bits if  $w_l = 1$  (since  $t$  is odd). Since every bit flip results in a change of the parity of the set of bits, the parity does not change if we flip an even number of bits and the parity also flips if we negate an odd number of bits. Therefore, if  $w_l$  is the correct (respectively wrong) parity of  $\{x_1, \dots, x_t\}$  then the parity of the added sets is 0 (respectively 1), i.e.,  $M'x' = 0^{2n}$  if  $Mx \oplus w = 0^{n/2}$  and  $M'x' = 1^{2n}$  if  $Mx \oplus w = 1^{n/2}$ . The number of ones in  $x' \in \{0, 1\}^{4n}$  is exactly  $2n$ . If  $t$  is even, we can just change the cases for the added edges such that we flip an even number of bits in the case  $w_l = 0$  and an odd number of bits if  $w_l = 1$ . Overall, this shows that a lower bound for  $BHH_{t,n}$  implies a lower bound for  $BHH_{t,4n}^0$ .  $\square$

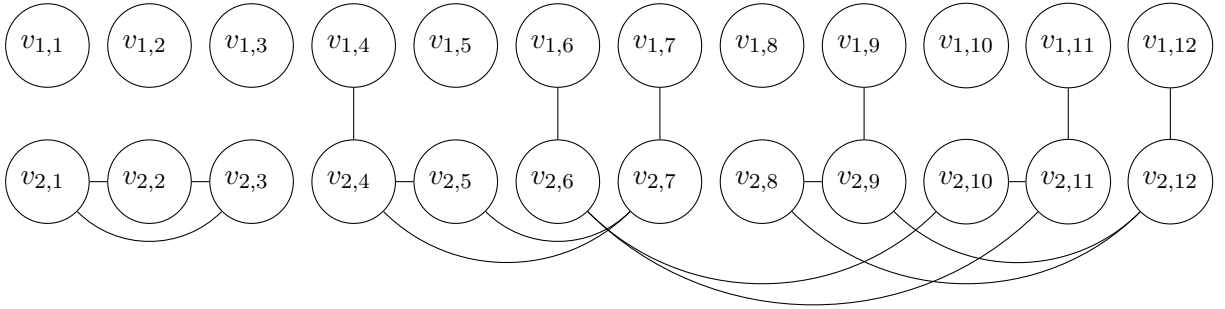


Figure 1: Worst case instance for  $t = 3$ . Bob's hypermatching corresponds to disjoint 3-cliques among the lower nodes and Alice's input vector corresponds to the edges between upper and lower nodes.

**Theorem 2.** *Any randomized streaming algorithm that approximates the maximum matching size within a  $1 + \frac{1}{3t/2-1}$  factor for  $t \geq 2$  needs  $\Omega(n^{1-1/t})$  space.*

*Proof.* Let  $x, M$  be the input to the  $BHH_{t,n}^0$  problem, i.e.,  $M$  is a perfect  $t$ -hypermatching on the coordinates of  $x$ ,  $x$  has exactly  $n/2$  ones and it is promised that either  $Mx = 0^{n/t}$  or  $Mx = 1^{n/t}$ . We construct the graph for the reduction as described above: For each bit  $x_i$  we have two nodes  $v_{1,i}, v_{2,i}$  and Alice adds the edge  $\{v_{1,i}, v_{2,i}\}$  iff  $x_i = 1$ . For each edge  $\{x_{i_1}, \dots, x_{i_t}\} \in M$  Bob adds a  $t$ -clique consisting of the nodes  $v_{2,i_1}, \dots, v_{2,i_t}$ . For now, let us assume  $t$  to be odd. We know that the matching is at least  $n/2$  because  $x$  has exactly  $n/2$  ones. Since Bob adds a clique for every edge it is always possible to match all (or all but one) nodes of the clique whose corresponding bit is 0. In the case of  $Mx = 0^{n/t}$  the parity of every edge is 0, i.e., the number of nodes whose corresponding bit is 1 is even. Let  $M_{2i} \subseteq M$  be the hyperedges containing exactly  $2i$  one bits and define  $l_{2i} := |M_{2i}|$ . Then we know  $n/2 = \sum_{i=0}^{\lfloor t/2 \rfloor} 2i \cdot l_{2i}$  and  $|M| = n/t = \sum_{i=0}^{\lfloor t/2 \rfloor} l_{2i}$ . For every edge in  $M_{2i}$  the size of the maximum matching within the corresponding subgraph is exactly  $2i + \lfloor (t - 2i)/2 \rfloor = 2i + \lfloor t/2 \rfloor - i$  for every  $i = 0, \dots, \lfloor t/2 \rfloor$  (see Fig. 1). Thus, we have a matching of size

$$\sum_{i=0}^{\lfloor t/2 \rfloor} (2i + (\lfloor t/2 \rfloor - i))l_{2i} = \frac{n}{2} + \frac{t-1}{2} \cdot \frac{n}{t} - \frac{n}{4} = \frac{3n}{4} - \frac{n}{2t}.$$

If we have  $Mx = 1^{n/t}$  then let  $M_{2i+1} \subseteq M$  be the hyperedges containing exactly  $2i+1$  one bits and define  $l_{2i+1} := |M_{2i+1}|$ . Again, we know  $n/2 = \sum_{i=0}^{\lfloor t/2 \rfloor} (2i+1) \cdot l_{2i+1}$  and  $|M| = n/t = \sum_{i=0}^{\lfloor t/2 \rfloor} l_{2i+1}$ . For every edge in  $M_{2i+1}$  the size of the maximum matching within the corresponding subgraph is exactly  $2i+1 + (t - 2i - 1)/2 = 2i+1 + \lfloor t/2 \rfloor - i$  for every  $i = 0, \dots, \lfloor t/2 \rfloor$ . Thus, the maximum matching has a size

$$\sum_{i=0}^{\lfloor t/2 \rfloor} (2i+1 + (\lfloor t/2 \rfloor - i))l_{2i+1} = \frac{n}{2} + \frac{t-1}{2} \cdot \frac{n}{t} - \frac{1}{2} \sum_{i=0}^{\lfloor t/2 \rfloor} (2i+1) \cdot l_{2i+1} + \frac{n}{2t} = \frac{3n}{4}.$$

For  $t$  even, the size of the matching is

$$\sum_{i=0}^{t/2} (2i + (t - 2i)/2)l_{2i} = \frac{n}{2} + \frac{t}{2} \cdot \frac{n}{t} - \frac{n}{4} = \frac{3n}{4}$$

if  $Mx = 0^{n/t}$ . Otherwise, we have

$$\begin{aligned} \sum_{i=0}^{t/2} \left( 2i+1 + \left\lfloor \frac{t-2i-1}{2} \right\rfloor \right) l_{2i+1} &= \frac{n}{2} + \sum_{i=0}^{t/2} (t/2 - i - 1)l_{2i+1} \\ &= \frac{n}{2} - (t/2 - 1) \cdot \frac{n}{t} - \frac{n}{4} + \frac{n}{2t} = \frac{3n}{4} - \frac{n}{2t}. \end{aligned}$$

As a consequence, every streaming algorithm that computes an  $\alpha$ -approximation on the size of a maximum matching with

$$\alpha < \frac{(3/4)n}{((3/4) - 1/(2t))n} = 1/(1 - 4/6t) = 1 + \frac{1}{3t/2 - 1}$$

can distinguish between  $Mx = 0^{n/t}$  and  $Mx = 1^{n/t}$  and, thus, needs  $\Omega(n^{1-1/t})$  space.  $\square$

Finally, constructing the Tutte-matrix with randomly chosen entries gives us

**Corollary 2.** *Any randomized streaming algorithm that approximates  $\text{rank}(A)$  of  $A \in \mathbb{R}^{n \times n}$  within a  $1 + \frac{1}{3t/2-1}$  factor for  $t \geq 2$  requires  $\Omega(n^{1-1/t})$  space.*

## References

- [1] K. Ahn and S. Guha. Graph sparsification in the semi-streaming model. In *ICALP (2)*, pages 328–338, 2009.
- [2] K. Ahn, S. Guha, and A. McGregor. Analyzing graph structure via linear measurements. In *SODA*, pages 459–467, 2012.
- [3] K. Ahn, S. Guha, and A. McGregor. Graph sketches: sparsification, spanners, and subgraphs. In *PODS*, pages 5–14, 2012.
- [4] K. Ahn, S. Guha, and A. McGregor. Spectral sparsification in dynamic graph streams. In *APPROX-RANDOM*, pages 1–10, 2013.
- [5] S. Assadi, S. Khanna, Y. Li, and G. Yaroslavtsev. Tight bounds for linear sketches of approximate matchings. *CoRR*, abs/1505.01467, 2015.
- [6] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. *SIAM J. Comput.*, 38(1):366–384, 2008.
- [7] N. Barkay, E. Porat, and B. Shalem. Efficient sampling of non-strict turnstile data streams. In *FCT*, pages 48–59, 2013.
- [8] R. Chitnis, G. Cormode, H. Esfandiari, M. Hajiaghayi, A. McGregor, M. Monemizadeh, and S. Vorotnikova. Kernelization via sampling with applications to dynamic graph streams. *CoRR*, abs/1505.01731, 2015.
- [9] K. Clarkson and D. Woodruff. Numerical linear algebra in the streaming model. In *STOC*, pages 205–214, 2009.
- [10] M. Crouch, A. McGregor, and D. Stubbs. Dynamic graphs in the sliding-window model. In *ESA*, pages 337–348, 2013.
- [11] M. Crouch and D. Stubbs. Improved streaming algorithms for weighted matching, via unweighted matching. In *APPROX/RANDOM*, pages 96–104, 2014.
- [12] A. Czygrinow, M. Hanckowiak, and E. Szymanska. Fast distributed approximation algorithm for the maximum matching problem in bounded arboricity graphs. In *ISAAC*, pages 668–678, 2009.
- [13] L. Epstein, A. Levin, J. Mestre, and D. Segev. Improved approximation guarantees for weighted matching in the semi-streaming model. *SIAM J. Discrete Math.*, 25(3):1251–1265, 2011.
- [14] L. Epstein, A. Levin, D. Segev, and O. Weimann. Improved bounds for online preemptive matching. In *STACS*, pages 389–399, 2013.
- [15] H. Esfandiari, M. Hajiaghayi, V. Liaghat, M. Monemizadeh, and K. Onak. Streaming algorithms for estimating the matching size in planar graphs and beyond. In *SODA*, pages 1217–1233, 2015.
- [16] J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang. On graph problems in a semi-streaming model. *Theor. Comput. Sci.*, 348(2-3):207–216, 2005.
- [17] D. Gavinsky, J. Kempe, I. Kerenidis, R. Raz, and R. de Wolf. Exponential separation for one-way quantum communication complexity, with applications to cryptography. *SIAM J. Comput.*, 38(5):1695–1708, 2008.

- [18] A. Goel, M. Kapralov, and S. Khanna. On the communication and streaming complexity of maximum bipartite matching. In *SODA*, pages 468–485, 2012.
- [19] M. Henzinger, P. Raghavan, and S. Rajagopalan. Computing on data streams, 1998.
- [20] H. Jowhari, M. Saglam, and G. Tardos. Tight bounds for lp samplers, finding duplicates in streams, and related problems. In *PODS*, pages 49–58, 2011.
- [21] D. Kane, J. Nelson, and D. Woodruff. An optimal algorithm for the distinct elements problem. In *PODS*, pages 41–52, 2010.
- [22] M. Kapralov. Better bounds for matchings in the streaming model. In *SODA*, pages 1679–1697, 2013.
- [23] M. Kapralov, S. Khanna, and M. Sudan. Approximating matching size from random streams. In *SODA*, pages 734–751, 2014.
- [24] C. Konrad. Maximum matching in turnstile streams. *CoRR*, abs/1505.01460, 2015.
- [25] C. Konrad, F. Magniez, and C. Mathieu. Maximum matching in semi-streaming with few passes. In *APPROX-RANDOM*, pages 231–242, 2012.
- [26] Y. Li, H. Nguyen, and D. Woodruff. On sketching matrix norms and the top singular vector. In *SODA*, pages 1562–1581, 2014.
- [27] Y. Li, H. Nguyen, and D. Woodruff. Turnstile streaming algorithms might as well be linear sketches. In *STOC*, pages 174–183, 2014.
- [28] L. Lovász. On determinants, matchings, and random algorithms. In *FCT*, pages 565–574, 1979.
- [29] A. McGregor. Finding graph matchings in data streams. In *APPROX-RANDOM*, pages 170–181, 2005.
- [30] A. McGregor. Graph stream algorithms: a survey. *SIGMOD Record*, 43(1):9–20, 2014.
- [31] W. Tutte. The factorization of linear graphs. *J. London Math. Soc.*, 22:107–111, 1947.
- [32] R. Uehara and Z. Chen. Parallel approximation algorithms for maximum weighted matching in general graphs. *Inf. Process. Lett.*, 76(1-2):13–17, 2000.
- [33] E. Verbin and W. Yu. The streaming complexity of cycle counting, sorting by reversals, and other problems. In *SODA*, pages 11–25. SIAM, 2011.
- [34] M. Zelke. Weighted matching in the semi-streaming model. *Algorithmica*, 62(1-2):1–20, 2012.